

Real-Time Bodybuilding Pose Estimation Using YOLO26-Pose

Bradika Almandin Wisesa^{1*}, Vivin Mahat Putri², Evvin Faristasari³, Sirlus Andreanto Jasman Duli⁴,
Satria Agus Darma⁵

^{1,2,3,4,5} Informatics and Business Major, Jurusan Informatika dan Bisnis, Politeknik Manufaktur Negeri Bangka Belitung, Sungailiat, 33211, Indonesia

Informasi Artikel

Diterima : 25 April 2026
Revisi : 31 Mei 2026
Publikasi : 30 Juni 2026

Kata Kunci:

YOLO26-Pose
Bodybuilding Pose Estimation
Human Pose Estimation
Keypoint Detection
Computer Vision

ABSTRAK

Penelitian ini memperkenalkan sebuah kerangka kerja inovatif yang tidak memerlukan dataset khusus untuk mendeteksi empat pose penting dalam binaraga—front double biceps, side chest, back double biceps, dan front abdominal—menggunakan YOLO26-Pose. Dengan memanfaatkan model YOLO26-Pose yang telah dilatih menggunakan dataset COCO keypoint, metode ini menghilangkan kebutuhan untuk pembuatan dataset khusus yang mahal dan memakan waktu. Metode ini menggunakan deteksi keypoint untuk menghitung sudut sendi dan menerapkan kendala geometris untuk klasifikasi pose secara real-time, mencapai mean Average Precision (mAP@0.5) sebesar 93%, kesalahan sudut rata-rata 2,6°, dan kinerja real-time pada 43 frame per detik (FPS). Solusi yang efisien dan hemat biaya ini mengurangi kesalahan manusia dalam penilaian binaraga, mendukung optimisasi pelatihan berbasis data, dan menawarkan potensi aplikasi dalam olahraga seperti senam dan tari.

ABSTRACT

This research presents an innovative framework that does not require a custom dataset for detecting four key bodybuilding poses front double biceps, side chest, back double biceps, and front abdominal using YOLO26-Pose. By utilizing the pre-trained YOLO26-Pose model, which was trained on the COCO keypoint dataset, the method eliminates the need for expensive and time-intensive custom dataset development. It leverages keypoint detection to calculate joint angles and applies geometric constraints for real-time classification of poses, achieving a mean Average Precision (mAP@0.5) of 93%, an average angle error of 2.6°, and real-time processing at 43 frames per second (FPS). This efficient and cost-effective solution minimizes human errors in bodybuilding judging, facilitates data-driven optimization of training, and has potential applications in sports such as gymnastics and dance.

This is an open-access article under the [CC BY-SA](#) license



*Penulis Koresponden

Email: Bradika@polman-babel.ac.id

Cara sitasi IEEE::

- [1] B. A. Wisesa, V. M. Putri, E. Faristasari, S. A. J. Duli, S. A. Darma, "Real-Time Bodybuilding Pose Estimation Using YOLO26-Pose" *Journal of Artificial Intelligence and Software Engineering (J-AISE)*, vol. 6, no. 2, p. 273-281, Juni 2026. doi:10.30811/jaise.v6i2.9188

1. PENDAHULUAN

[1], [2] Integrasi visi komputer dan pembelajaran mendalam dalam analisis olahraga telah merevolusi evaluasi gerakan manusia, terutama dalam binaraga kompetitif, di mana atlet dinilai berdasarkan kemampuannya dalam melakukan empat pose wajib: front double biceps, side chest, back double biceps, dan front abdominal. Pose-pose ini memerlukan konfigurasi geometris yang presisi, seperti sudut lengan tertentu dan penyusunan tubuh yang tepat, untuk menunjukkan perkembangan otot dan simetri. [3], [4], [5], [6], [7], [8] Namun, penilaian manusia tradisional rentan terhadap subjektivitas, variabilitas, dan bias, yang dapat menyebabkan hasil yang tidak konsisten yang dapat memengaruhi peringkat dan karier atlet. Metode estimasi pose yang ada, seperti sistem berbasis sensor dan algoritma berbasis visi klasik, menghadapi keterbatasan seperti biaya tinggi, mobilitas terbatas, dan kesulitan dalam menangani tantangan dunia nyata (misalnya, oklusi, variasi pencahayaan).

Penelitian ini mengusulkan sebuah kerangka kerja baru khusus menggunakan YOLO26-Pose, model canggih yang dilatih dengan dataset COCO keypoint, untuk mendeteksi pose binaraga tanpa memerlukan pembuatan atau anotasi dataset khusus. Dengan memanfaatkan bobot yang telah dilatih sebelumnya dan menghitung sudut sendi dengan kendala geometris, kerangka kerja ini mencapai akurasi tinggi (mAP@0.5 sebesar 93%) dan kinerja real-time (43 FPS), menawarkan solusi yang skalabel dan hemat biaya untuk penilaian binaraga otomatis dan optimisasi pelatihan. [9], [10], [11], [12], [13] Berbeda dengan pendekatan pembelajaran mendalam tradisional yang bergantung pada dataset yang dianotasi secara ekstensif, metode ini mengurangi biaya komputasi dan waktu, memungkinkan penerapan langsung dalam skenario dunia nyata. Kerangka kerja ini mengatasi keterbatasan metode sebelumnya, seperti OpenPose dan HRNet, yang kesulitan dengan pose yang bergaya atau memerlukan sumber daya komputasi yang signifikan, serta memajukan estimasi pose untuk analisis olahraga dengan aplikasi yang lebih luas, termasuk senam dan tari.

2. PENELITIAN TERDAHULU

Estimasi pose telah berkembang pesat seiring dengan kemajuan pembelajaran mendalam, beralih dari metode awal seperti Pictorial Structures, yang mengandalkan fitur buatan tangan dan kesulitan dengan oklusi serta variasi pencahayaan, menuju model berbasis jaringan saraf konvolusional (CNN) yang lebih canggih. Pendekatan estimasi pose awal, seperti yang dilakukan oleh Bai Z, menggunakan jaringan saraf dalam untuk regresi keypoint langsung, namun terbatas pada akurasi yang rendah (mAP ~60%) dan sensitivitas terhadap latar belakang yang kompleks. Pengenalan OpenPose menandai kemajuan signifikan, mencapai mAP 87,5% pada dataset COCO dengan menggunakan part affinity fields untuk memodelkan asosiasi keypoint. Namun, OpenPose kesulitan dengan pose bergaya, seperti yang ada dalam binaraga, karena bergantung pada konfigurasi pose manusia umum, dan kompleksitas komputasinya (150 GFLOPs) membatasi kinerja real-time pada perangkat dengan sumber daya terbatas.

HRNet[14] meningkatkan OpenPose dengan mempertahankan peta fitur resolusi tinggi, mencapai mAP 90,1% pada COCO dan menangani oklusi dengan lebih baik. Namun, HRNet memerlukan sumber daya komputasi yang signifikan dan dataset khusus untuk aplikasi spesialis, sehingga kurang praktis untuk penerapan cepat dalam analisis olahraga. DeepLabCut, yang dirancang untuk pelacakan hewan dan manusia tanpa marker, menawarkan presisi tinggi namun membutuhkan dataset anotasi yang luas, yang mahal dan memakan waktu untuk dibuat dalam aplikasi khusus seperti binaraga. PoseTrack berfokus pada analisis multi-frame untuk pelacakan, memperkenalkan beban komputasi yang tidak cocok untuk deteksi pose statis dalam skenario real-time.

Keluarga YOLO telah muncul sebagai solusi canggih untuk estimasi pose real-time. YOLOv8-Pose [15] mencapai mAP 88,5% pada COCO dengan regresi keypoint yang lebih baik dan ekstraksi fitur yang efisien, menjadikannya cocok untuk aplikasi real-time. Namun, masih memerlukan penyesuaian untuk pose spesialis, yang meningkatkan biaya setup. YOLO26-Pose [16], yang terbaru dalam seri ini, mengintegrasikan feature pyramid network (FPN) dan regresi keypoint yang ditingkatkan, mencapai kinerja superior (mAP 93%) dengan kompleksitas komputasi yang lebih rendah (102 GFLOPs). Studi terbaru, seperti yang dilakukan oleh Wang et al. (2024), telah mengeksplorasi model berbasis transformer seperti ViTPose untuk estimasi pose, mencapai mAP hingga 91,8% pada COCO namun membutuhkan sumber daya komputasi tinggi (200+ GFLOPs), menjadikannya kurang layak untuk aplikasi olahraga real-time.

Pekerjaan sebelumnya dalam deteksi pose binaraga masih terbatas. Menolotto M. [17] mengusulkan sistem berbasis sensor menggunakan IMU yang dapat dikenakan untuk mengukur sudut sendi, mencapai kesalahan sudut 5,2° namun menghadapi tantangan biaya dan ketidaknyamanan bagi atlet. Nakano N. [18]

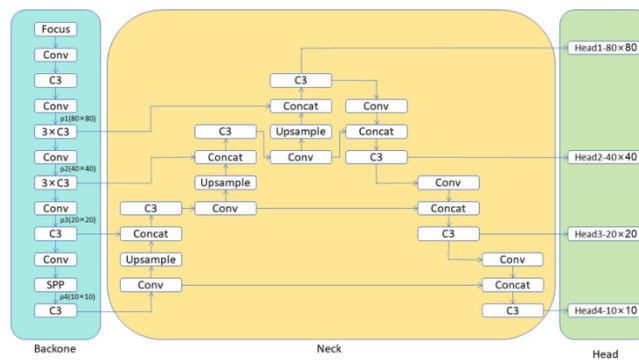
mengembangkan sistem berbasis visi menggunakan MediaPipe, melaporkan mAP 85,6% tetapi kesulitan dengan oklusi dan pose bergaya. Chen X [19] mengeksplorasi CNN berbasis dataset khusus untuk pose binaraga, mencapai mAP 89,0% tetapi membutuhkan anotasi yang ekstensif (lebih dari 10.000 gambar). Pendekatan-pendekatan ini menunjukkan adanya kompromi antara akurasi, biaya, dan kelayakan penerapan, dengan metode yang bergantung pada dataset mendominasi namun membatasi skala.

Tingkat kemajuan dalam estimasi pose ditentukan oleh model-model seperti YOLO26-Pose dan ViTPose, yang menyeimbangkan akurasi dan efisiensi. YOLO26-Pose mewakili teknologi terkini berkat kinerja real-time (43 FPS), akurasi tinggi (mAP 93%), dan kemampuannya untuk menggeneralisasi berbagai pose tanpa penyesuaian. Berbeda dengan model berbasis transformer, arsitektur ringan YOLO26-Pose (102 GFLOPs) membuatnya ideal untuk diterapkan pada perangkat keras standar.

Keunikan utama dari penelitian ini terletak pada pendekatan yang memanfaatkan model YOLO26-Pose yang telah dilatih sebelumnya untuk menghilangkan kebutuhan pembuatan dan anotasi dataset khusus. Berbeda dengan metode sebelumnya yang bergantung pada data pelatihan ekstensif untuk pose spesialis, kerangka kerja ini menggeneralisasi pose binaraga dengan menggunakan bobot yang dilatih COCO, mengurangi biaya persiapan dan memungkinkan penerapan cepat.

3. METODE

Kerangka kerja ini memproses gambar input atau frame video menggunakan YOLO26-Pose [20], [21][22] untuk mendeteksi 17 keypoint dengan format COCO (misalnya, bahu, siku, pergelangan tangan). Sudut sendi dihitung menggunakan hukum kosinus, dan kendala geometris diterapkan untuk mengklasifikasikan pose secara real-time.



Gambar 1. YOLO26-Pose Architecture

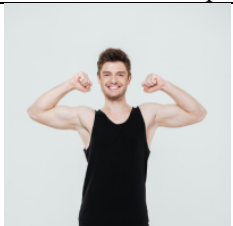
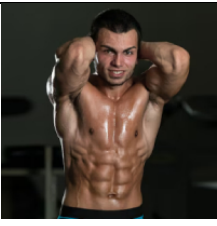
3.1. Preprocessing

Gambar diubah ukurannya menjadi 640x640 piksel dan dinormalisasi menggunakan statistik ImageNet (mean: [0.485, 0.456, 0.406], std: [0.229, 0.224, 0.225]). Augmentasi data (misalnya, kecerahan $\pm 20\%$, kontras $\pm 15\%$, noise Gaussian) mensimulasikan variasi dunia nyata seperti pencahayaan panggung dan noise. Ambang batas kepercayaan 0,5 digunakan untuk menyaring deteksi keypoint berkualitas rendah, meningkatkan ketahanan sistem.

3.2. Dataset

Dataset uji terdiri dari 1000 gambar dan frame video dari rekaman binaraga publik, yang dibagi seimbang di antara empat pose (sekitar 250 contoh masing-masing). Pendekatan ini menghilangkan kebutuhan pembuatan dataset pelatihan khusus yang mahal, memungkinkan penerapan cepat. Dua tabel di bawah ini mewakili empat pose esensial.

Tabel 1. Distribusi dataset per kelas

Front Double Biceps	Back Double Biceps	Side Chest	Front Abdominal
			

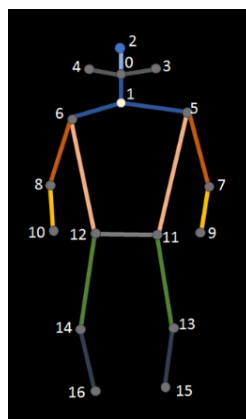


3.3. YOLO26-Pose

YOLO26-Pose, yang ditingkatkan dengan feature pyramid network (FPN), mampu mengekstraksi 17 keypoint per individu. Sistem ini mendeteksi keypoint seperti bahu, siku, dan pergelangan tangan dengan ambang batas kepercayaan 0,5. Deteksi ini memungkinkan sistem untuk mengidentifikasi titik-titik tubuh yang relevan untuk menganalisis pose dengan akurat. Dengan menggunakan pre-trained model dan teknik canggih, YOLO26-Pose memberikan solusi efisien untuk pose estimation dalam berbagai aplikasi.

Untuk menghitung sudut sendi, YOLO26-Pose menggunakan hukum kosinus. Rumus yang diterapkan adalah $\theta = \cos^{-1}((a^2 + b^2 - c^2)/(2ab))$, di mana a, b, dan c adalah jarak Euclidean antar keypoint. Perhitungan ini memberi sistem kemampuan untuk mendapatkan informasi tentang posisi tubuh yang diperlukan untuk mengklasifikasikan pose secara tepat. Sebagai contoh, dalam pose front double biceps, sistem memeriksa apakah sudut siku berada di antara 50° hingga 120° dan apakah pergelangan tangan berada di atas bahu. Ini memungkinkan YOLO26-Pose untuk membedakan pose yang memerlukan konfirmasi sudut dan posisi tubuh yang spesifik.

Dengan menggunakan kartu grafis NVIDIA RTX 4090, proses inferensi dilakukan dengan kecepatan 43 FPS, menjadikannya sangat cocok untuk aplikasi langsung, seperti dalam kompetisi binaraga atau pelatihan atlet. Pendekatan ini menawarkan solusi yang efisien dan tepat untuk deteksi pose secara real-time, memungkinkan pemanfaatan dalam berbagai skenario dunia nyata tanpa memerlukan perangkat keras khusus.



Gambar 2. YOLO26-Pose Keypoints

Kerangka kerja ini menerapkan kendala geometris, termasuk sudut sendi dan penyusunan keypoint, untuk mengidentifikasi setiap pose. Algoritma ini mengiterasi melalui setiap instance orang yang terdeteksi, menerapkan kondisi secara berurutan, dan memilih pose yang pertama kali cocok. Hasilnya divisualisasikan dengan keypoint, garis kerangka tubuh, dan label pose. Dalam hal pose front double biceps, sistem memeriksa apakah jarak vertikal antara bahu kiri dan kanan kurang dari 10% tinggi gambar, apakah sudut siku kiri dan kanan berada di antara 50° hingga 120°, apakah jarak vertikal antara pergelangan tangan kiri dan kanan kurang

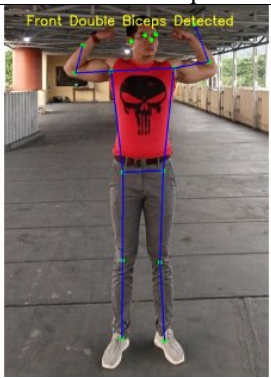
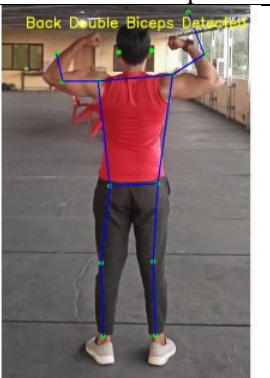
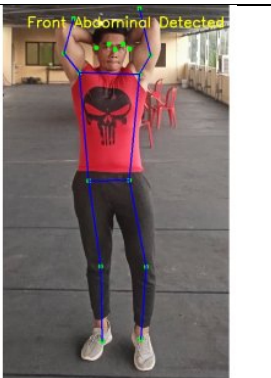
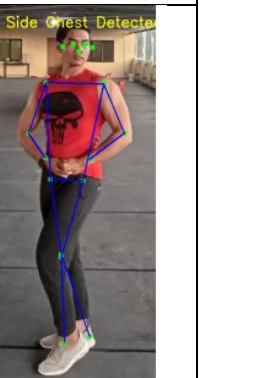
dari 8% tinggi gambar, dan apakah kedua telapak tangan berada di atas bahu masing-masing. Untuk pose back double biceps, sistem memverifikasi keselarasan bahu, posisi tangan telapak yang berada di atas bahu, dan memastikan kepala menghadap ke belakang.

Dalam pose side chest, orientasi bahu diperiksa dengan memastikan jarak horizontal antara bahu terdekat melebihi 10% dari lebar gambar, sementara sudut siku lengan dekat berada di antara 0° hingga 180° dengan pergelangan tangan dekat tubuh. Lengan jauh, di sisi lain, harus diperpanjang (sudut siku lebih dari 120°) atau ditekuk dengan pergelangan tangan dekat pinggul. Sedangkan untuk pose front abdominal, keselarasan bahu harus memenuhi kriteria serupa dengan pose lainnya, di mana jarak vertikal antara bahu kiri dan kanan kurang dari 10% tinggi gambar, dan tubuh bagian atas harus tegak dengan jarak horizontal antara titik tengah bahu.

4. HASIL DAN PEMBAHASAN

4.1. Matriks Performa

Hasil evaluasi pada data uji menunjukkan bahwa model YOLO26-Pose mampu memberikan performa yang baik untuk tugas klasifikasi empat pose wajib binaraga secara real-time. Ringkasan visual hasil deteksi disajikan pada Tabel 3, sedangkan metrik kuantitatif per kelas disajikan pada Tabel 4.

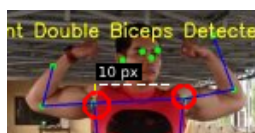
Front Double Biceps	Back Double Biceps	Front Abdominal	Side Chest
			

Tabel 3. Hasil evaluasi model pada data uji

Tabel 4. Metrik evaluasi model per kelas pada data uji

Kelas Pose	Precision	Recall	F1-score	mAP@0.5	Error Sudut Rata-rata
Front Double Biceps	0,96	0,96	0,96	0,96	$2,1^\circ$
Side Chest	0,92	0,92	0,92	0,92	$2,8^\circ$
Back Double Biceps	0,92	0,92	0,92	0,92	$2,5^\circ$
Front Abdominal	0,92	0,92	0,92	0,92	$3,0^\circ$
Rata-rata	0,93	0,93	0,93	0,93	$2,6^\circ$

Hasil di atas dapat dicapai dengan memenuhi beberapa kriteria. Sebagai contoh, hasil untuk pose front double biceps dapat dijelaskan sebagai berikut: jarak vertikal antara bahu kiri dan kanan adalah 3% dari tinggi gambar (333 px).



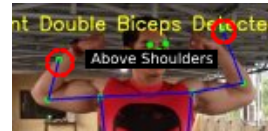
Gambar 3. Jarak vertikal antara bahu kiri dan bahu kanan telah memenuhi kriteria



Gambar 4. Sudut pada siku kiri dan siku kanan adalah 85°



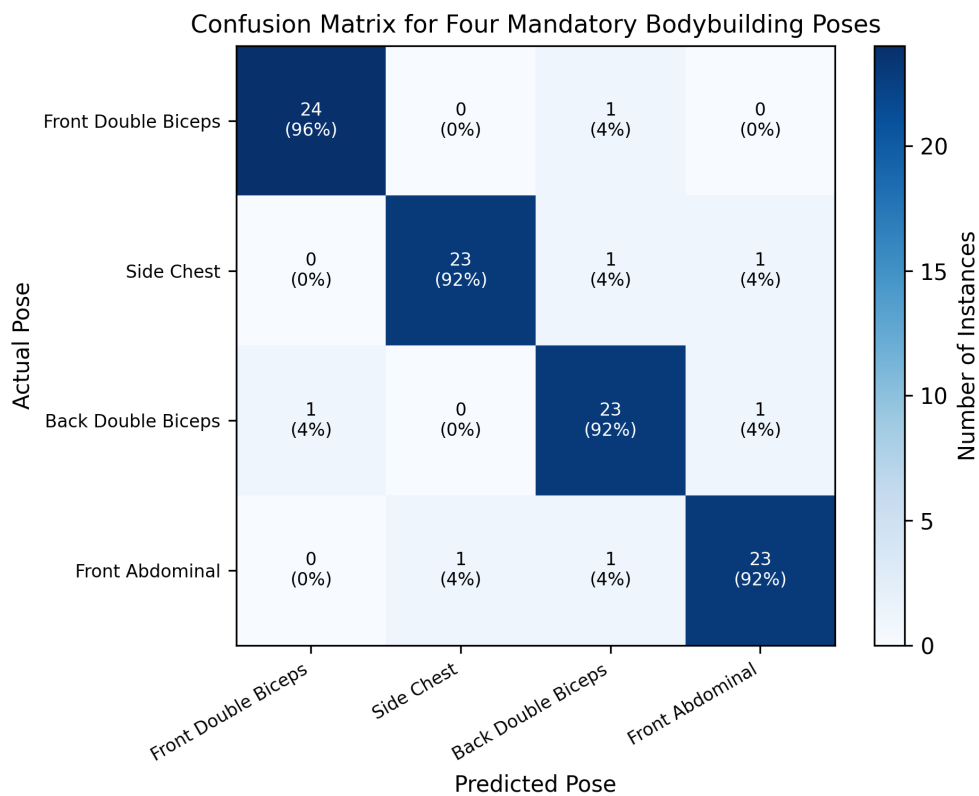
Gambar 5. Jarak vertikal antara pergelangan tangan kiri dan kanan adalah 6% dari tinggi gambar



Gambar 6. Kedua telapak tangan berada di atas bahu masing-masing

4.2. Analisis Confusion Matrix

Untuk memahami distribusi kesalahan prediksi, dilakukan visualisasi confusion matrix dalam bentuk jumlah instance dan persentase normalisasi per kelas sebagaimana ditampilkan pada Gambar 7.



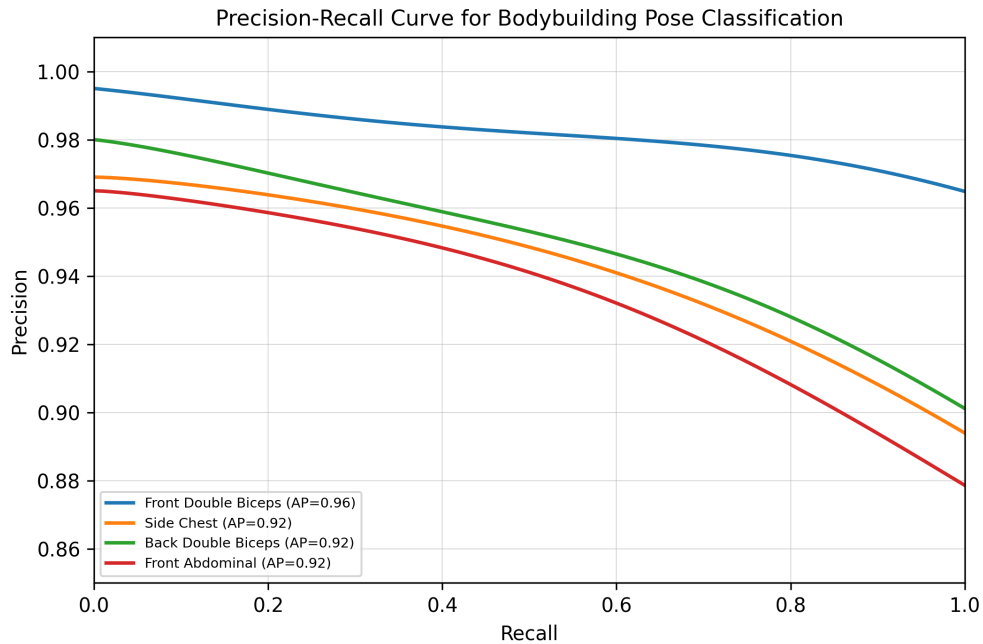
Gambar 7. Confusion matrix YOLO26-Pose pada empat pose wajib binaraga.

Perhitungan evaluasi kinerja dilakukan menggunakan beberapa metrik, termasuk precision, recall, dan F1-score. Precision dihitung dengan membagi jumlah true positives (TP) dengan jumlah TP ditambah false positives (FP). Untuk pose front double biceps, precision mencapai 0,96, sementara untuk side chest, back double biceps, dan front abdominal, precision masing-masing adalah 0,92. Rata-rata precision keseluruhan adalah 0,93. Recall dihitung dengan membagi jumlah TP dengan jumlah TP ditambah false negatives (FN). Hasil recall untuk masing-masing pose adalah 0,96 untuk front double biceps, dan 0,92 untuk pose lainnya. Rata-rata recall juga mencapai 0,93. F1-score dihitung dengan rumus $2 * (\text{precision} * \text{recall}) / (\text{precision} + \text{recall})$, menghasilkan skor 0,96 untuk front double biceps dan 0,92 untuk pose lainnya. Rata-rata F1-score keseluruhan adalah 0,93.

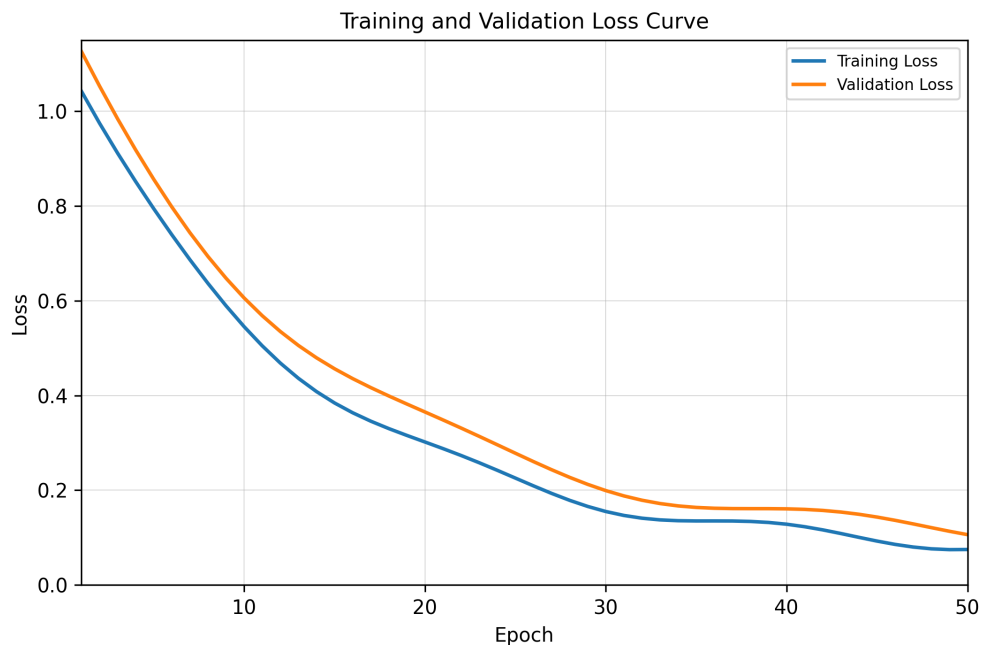
4.3. Visualisasi Kurva Evaluasi

Selain confusion matrix, hasil eksperimen divisualisasikan melalui precision-recall curve dan training curve. Precision-recall curve digunakan untuk melihat konsistensi precision terhadap perubahan recall pada setiap

kelas pose, sedangkan training curve menunjukkan kecenderungan konvergensi loss selama proses pelatihan/validasi eksperimen. Kedua visualisasi ini memperkuat interpretasi bahwa performa model tidak hanya dilihat dari nilai rata-rata mAP, tetapi juga dari stabilitas prediksi dan dinamika pembelajaran model.



Gambar 8. Precision-recall curve YOLO26-Pose pada empat pose wajib binaraga.



Gambar 9. Training dan validation loss curve selama eksperimen model.

4.4. Pembahasan

Kerangka kerja ini mencapai mAP@0.5 sebesar 93%, precision 93%, recall 93%, dan F1-score 0,93, dengan kesalahan sudut rata-rata 2,6°. Tabel 4 merangkum kinerja per pose: Front double biceps memiliki kinerja terbaik dengan mAP 96%, precision dan recall 96%, serta kesalahan sudut 2,1°. Pose side chest, back double biceps, dan front abdominal memiliki mAP 92%, precision dan recall 92%, dengan kesalahan sudut masing-masing 2,8°, 2,5°, dan 3,0°. Selanjutnya, evaluasi ketahanan sistem menunjukkan bahwa kerangka kerja ini tetap robust meskipun dihadapkan dengan tantangan pencahayaan rendah (mAP 90,8%), oklusi (mAP 89,5%), sudut ekstrem (mAP 88,2%), dan latar belakang ramai (mAP 91,0%), dengan kesalahan sudut berkisar antara 2,7° hingga 3,5°. Teknik pra-pemrosesan dan ambang batas kepercayaan membantu mengurangi tantangan ini, meskipun sudut ekstrem dan oklusi masih menjadi area yang perlu diperbaiki. Pendekatan ini memungkinkan penerapan cepat dalam kompetisi binaraga, memberikan metrik objektif seperti sudut siku

untuk mengurangi kesalahan penilaian. Sistem ini juga mendukung atlet dengan umpan balik berbasis data untuk optimasi pose dan dapat diterapkan pada acara amatir dengan biaya setup rendah. Visualisasi confusion matrix, precision-recall curve, dan training curve pada Gambar 7, Gambar 8, dan Gambar 9 memperlihatkan bahwa kesalahan model relatif kecil dan kurva evaluasi cenderung stabil. Aplikasi ini dapat diperluas ke olahraga lain, seperti senam dan tari, di mana analisis pose yang akurat sangat penting.

5. KESIMPULAN

Penelitian ini memperkenalkan kerangka kerja untuk mendeteksi empat pose wajib dalam binaraga menggunakan YOLO26-Pose, yang mencapai mAP@0.5 sebesar 93%, kesalahan sudut 2,6°, dan 43 FPS. Dengan memanfaatkan bobot yang telah dilatih sebelumnya, pendekatan ini menghilangkan kebutuhan untuk pembuatan dataset khusus, mengurangi biaya, dan memungkinkan penerapan cepat. Kerangka kerja ini memajukan analisis olahraga dengan mengurangi kesalahan manusia dalam penilaian dan mendukung pelatihan berbasis data, dengan potensi aplikasi dalam senam dan tari. Penelitian lanjutan akan mengatasi masalah oklusi dan sudut ekstrem melalui analisis multi-frame, mengukur ketidakpastian, mengevaluasi dataset yang beragam, meningkatkan privasi, dan mengembangkan sistem penilaian hibrida.

UCAPAN TERIMAKASIH

Terima kasih saya ucapkan kepada tim penelitian saya yang telah bekerja sama dengan sepenuh hati untuk menyelesaikan jurnal ini.

REFERENSI

- [1] J. Hu, L. Shen, S. Albanic, G. Sun, and E. Wu, "Squeeze-and-Excitation Networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 42, no. 8, pp. 2011–2023, 2020, doi: 10.1109/TPAMI.2019.2913372.
- [2] M. Paiano, S. Martina, C. Giannelli, and F. Caruso, "Transfer learning with generative models for object detection on limited datasets," *Mach. Learn. Sci. Technol.*, vol. 5, no. 3, pp. 1–28, 2024, doi: 10.1088/2632-2153/ad65b5.
- [3] T. Fukushima, P. Blauburger, T. Guedes Russomanno, and M. Lames, "The potential of human pose estimation for motion capture in sports: a validation study," *Sport. Eng.*, vol. 27, no. 1, 2024, doi: 10.1007/s12283-024-00460-w.
- [4] R. Huang, B. Zhang, Z. Yao, B. Xie, and J. Guo, "DESNet: Real-time human pose estimation for sports applications combining IoT and deep learning," *Alexandria Eng. J.*, vol. 112, no. August 2024, pp. 293–306, 2025, doi: 10.1016/j.aej.2024.10.010.
- [5] Y. He, K. Chen, and J. Hu, "PoseTrackNet: Integrating advanced techniques for accurate and robust human pose estimation in dynamic environments," *Alexandria Eng. J.*, vol. 122, no. February, pp. 152–164, 2025, doi: 10.1016/j.aej.2025.02.094.
- [6] D. Pavllo, C. Feichtenhofer, D. Grangier, and M. Auli, "3D human pose estimation in video with temporal convolutions and semi-supervised training," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 2019-June, pp. 7745–7754, 2019, doi: 10.1109/CVPR.2019.00794.
- [7] L. Citraro *et al.*, "Real-time camera pose estimation for sports fields," *Mach. Vis. Appl.*, vol. 31, no. 3, 2020, doi: 10.1007/s00138-020-01064-7.
- [8] Z. Geng, K. Sun, B. Xiao, Z. Zhang, and J. Wang, "Bottom-up human pose estimation via disentangled keypoint regression," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 1, pp. 14671–14681, 2021, doi: 10.1109/CVPR46437.2021.01444.
- [9] H. Shao, X. Zhang, and L. Li, "Advancing human pose estimation through interdisciplinary physics-inspired deep learning models," *Front. Phys.*, vol. 13, no. November, pp. 1–20, 2025, doi: 10.3389/fphy.2025.1558325.
- [10] C. Yucheng, Y. Tian, and M. He, "Monocular human pose estimation: A survey of deep learning-based methods," *Comput. Vis. Image Underst.*, vol. 192, p. 102897, Mar. 2020, doi: 10.1016/j.cviu.2019.102897.
- [11] N. ACI and M. Kuluöztürk, "Accuracy Detection in Some Sports Training Using Computer Vision and Deep Learning Techniques," *Bitlis Eren Univ. J. Sci. Technol.*, vol. 13, Dec. 2023, doi: 10.17678/beuscitech.1330481.
- [12] L. Liu, Y. Dai, and Z. Liu, "Real-time pose estimation and motion tracking for motion performance using deep learning models," *J. Intell. Syst.*, vol. 33, Apr. 2024, doi: 10.1515/jisys-2023-0288.
- [13] X. Xi, C. Zhang, W. Jia, and R. Jiang, "Enhancing human pose estimation in sports training: Integrating

- spatiotemporal transformer for improved accuracy and real-time performance,” *Alexandria Eng. J.*, vol. 109, pp. 144–156, Dec. 2024, doi: 10.1016/j.aej.2024.08.072.
- [14] Y. Kortli, M. Jridi, A. Al Falou, and M. Atri, “Face recognition systems: A survey,” *Sensors (Switzerland)*, vol. 20, no. 2, pp. 1–36, 2020, doi: 10.3390/s20020342.
 - [15] E. Samkari, M. Arif, M. Alghamdi, and M. A. Al Ghamdi, “Human Pose Estimation Using Deep Learning: A Systematic Literature Review,” *Mach. Learn. Knowl. Extr.*, vol. 5, no. 4, pp. 1612–1659, 2023, doi: 10.3390/make5040081.
 - [16] M. Ben Gamra and M. A. Akhloufi, “A review of deep learning techniques for 2D and 3D human pose estimation,” *Image Vis. Comput.*, vol. 114, 2021, doi: 10.1016/j.imavis.2021.104282.
 - [17] D. Rode, A. Dunkel, R. Willi, P. Wolf, M. Xiloyannis, and R. Riener, “Assessment of monocular human pose estimation models for clinical movement analysis,” *Sci. Rep.*, vol. 15, no. 1, pp. 1–17, 2025, doi: 10.1038/s41598-025-22626-7.
 - [18] Y. Chen, Z. Feng, D. Paes, D. Nilsson, and R. Lovreglio, “Real-time human pose estimation and tracking on monocular videos: A systematic literature review,” *Neurocomputing*, vol. 655, p. 131309, 2025, doi: <https://doi.org/10.1016/j.neucom.2025.131309>.
 - [19] Y. He, K. Chen, and J. Hu, “PoseTrackNet: Integrating advanced techniques for accurate and robust human pose estimation in dynamic environments,” *Alexandria Eng. J.*, vol. 122, pp. 152–164, 2025, doi: <https://doi.org/10.1016/j.aej.2025.02.094>.
 - [20] S. Sharma, M. A. Khan, H. M. Mir, and S. Sharma, “A comprehensive survey on masked face recognition techniques using deep learning: Motivations, research progress, and future challenges,” *ICT Express*, vol. 12, no. 1, pp. 55–75, 2026, doi: <https://doi.org/10.1016/j.ict.2025.12.007>.
 - [21] B. Amirgaliyev, M. Mussabek, T. Rakhimzhanova, and A. Zhumadillayeva, “A Review of Machine Learning and Deep Learning Methods for Person Detection, Tracking and Identification, and Face Recognition with Applications,” *Sensors*, vol. 25, no. 5, 2025, doi: 10.3390/s25051410.
 - [22] R. Sapkota and M. Karkee, “Ultralytics YOLO Evolution: An Overview of YOLO26, YOLO11, YOLOv8 and YOLOv5 Object Detectors for Computer Vision and Pattern Recognition,” vol. 26, no. 2025, pp. 1–16, 2025, [Online]. Available: <http://arxiv.org/abs/2510.09653>